AN EFFICIENT JPEG STEGANOGRAPHIC SCHEME BASED ON THE BLOCK ENTROPY OF DCT COEFFICIENTS

Chang Wang, Jiangqun Ni, *Member, IEEE*

School of Information Science and Technology, Sun Yat-sen University, Guangzhou, 510006, P.R. China

ABSTRACT

Steganography is the art of covert communication. This paper presents an efficient JPEG steganography scheme based on the block entropy of DCT coefficients and syndrome trellis coding (STC). The proposed cost function explores both the block complexity and distortion effects due to flipping and rounding errors. The STC provides multiple solutions to embed messages to a block of coefficients. The proposed scheme determines the best one with minimal distortion effect. In this way, the total distortions are significantly reduced, which corresponds to less detectability of steganalysis. Compared with similar schemes, experiment results demonstrate the superior performance of the proposed scheme in terms of secure embedding capacity against steganalysis.

***Index Terms*—** steganography, block entropy, syndrome-trellis code (STC), JPEG

1 INTRODUCTION

Steganography is the art of covert communication, in which the sender embeds secret message into the LSB plane of original image (cover) with shared key, and obtain the stego image. To conceal the very existence of communication, the stego images have to be statistically undetectable from cover images [1].

As a widely adopted format for image storage and transmission, JPEG steganography has become the domain of extensive research. And there are a lot of schemes developed for JPEG steganography, such as Jsteg [5], F5[3], Outguess [6] and MB [4]

Historically, Jsteg is one of the first steganographic method for JPEG image. This method embeds data by LSB substitution of quantized DCT coefficients. However， Jsteg can be easily detected by common steganalysis tools, such as PEV274 [2], due to the histogram distortion of DCT coefficients introduced in embedding.

F5 Algorithm and its variants [7-9] improved the security performance by increasing its embedding efficiency through matrix encoding. Later, Kim et.al [8] proposed an improved scheme of matrix encoding (MME) by only modifying the coefficients with less distortion.

Sachnev et.al. in [11] use a fast BCH syndrome coding to provide multiple codewords, where the one with minimum distortion is determined for data hiding. Significant improvement in security performance against steganalysis is achieved. Note that both the MME and the method in [11] require the original BMP image for data hiding.

In [1], T.Filler et.al. proposed a practical approach to minimizing embedding impact in steganography based on syndrome-trellis coding (STC), which can asymptotically achieve the performance bound. The STC framework employs an additive distortion function described by the set of local costs ρ_i, without having to share them with the receiver. The design of cost function opens up the possibility to guide the steganographic system to embed data in image region of "hard-to-predict", which leads to less detectability of steganalysis. It is the general consensus that further substantial increase in secure payload for steganography can be achieved by properly designing the cost function instead of improving the coding scheme.

the block entropy of DCT coefficients are adopted to evaluate the local complexity of JPEG image and then used to design the cost function for JPEG steganography. By incorporating the new cost function with the STC framework, the proposed scheme achieves considerable performance improvement in terms of secure payload.

The remainder of the paper is organized as follows. After review the minimal-distortion embedding framework, the proposed JPEG steganographic scheme is presented in Section II. Experiment results and analysis are included in Section III. Finally concluding remarks and future work are given in Section IV.

2 THE PROPOSED SCHEME

2.1 Minimal distortion embedding framework

In steganography, the transmitter communicates with the receiver by hiding her messages in trusted media, such as digital images, so that it is hard to distinguish the stego media from the covers. The message is generally hidden (embedded) in the cover image by slightly modifying some individual elements of cover (LSBs of pixels, quantized DCT coefficients, etc.). The problem of minimizing the embedding impact for single-letter distortion is well formulated in [1]. Let the binary vector $\mathbf{x}=(x_1,x_2,\cdots,x_n),\mathbf{y}=(y_1,y_2,\cdots,y_n)\in(0,1)^n$ and $\mathbf{m}=(m_1,m_2,\cdots,m_k)\in(0,1)^k$ represent the cover, stego and message, respectively. The additive cost function is defined as

$$D(\mathbf{x},\mathbf{y})=\sum_{i=1}^{n}\rho_i(\mathbf{x},y_i) \tag{1}$$

where $\rho_i(\mathbf{x},y_i)$ is the cost of changing the ith cover element x_i to y_i.With the syndrome coding, the minimal distortion embedding framework is formulated as

$$Emb(\mathbf{x},\mathbf{m})=\arg\min_{\mathbf{y}\in C(\mathbf{m})}D(\mathbf{x},\mathbf{y}) \tag{2}$$

$$\text{and } H\mathbf{y}=\mathbf{m} \tag{3}$$

where H is parity-check matrix of the code C and $C(\mathbf{m})$ is the coset corresponding to syndrome $\mathbf{m}$.

2.2 The proposed distortion function based on block entropy of DCT coefficients

Information theory has found wide applications in communication and computer science. The concept of entropy in information theory is used to describe the uncertainty of information source, i.e.,the source with higher entropy is less of structure. Note the factor that most existing universal steganalysis algorithms predict the stego image based on the artifacts of feature vector which deviates from cover image. The feature vector for cover images, however, describes the statistics characteristics of natural image and is usually "predictable" with natural image model. The image regions with high complexity, such as texture and edges, correspond to the ones of less predictability. Therefore embed the message in image regions of "hard-to-predict" leads to steganographic scheme of less detectability.

Although the DCT transform itself shows only limited spatial resolution, the block DCT in JPEG standard exhibits sufficient spatial-frequency resolution. The block entropy of the quantized DCT coefficients in JPEG compressed domain can be used as the metric to evaluate the complexity of the corresponding 8×8 image block. The proposed cost function ρ_i takes into account both the block entropy and distortion effect due to flipping, and is defined as

$$\rho_i = \rho_{i\_ent} * \rho_{i\_f} \tag{4}$$

where ρ_{i_ent} and ρ_{i_f} denote the cost of block entropy and flipping distortion for i th element, respectively. With the new cost function, those elements in cover image with high block entropy and less flipping distortion would have the priority to be modified.

For JPEG steganography, all non-zero quantized AC coefficients in DCT domain are scrambled with key K and formed the cover sequence $\mathbf{x}$. Let x_i be the i th DCT coefficient in $\mathbf{x}$, and belongs to DCT block $B(i)$. Assume there are K different non-zero AC coefficients in block $B(i)$, and let $b(i)_k, k = 1,\cdots,K$ denote the k th AC coefficients in $B(i)$ with frequency of occurrence $p\left(b(i)_k\right)$. The block entropy for x_i is computed according to defined as follows

$$H\left(B(i)\right) = -\sum_{k=1}^{K} p\left(b(i)_k\right) \log p\left(b(i)_k\right) \tag{5}$$

And the block entropy cost for x_i is defined as

$$\rho_{i\_ent} = 1 \big/ H\left(B(i)\right)^{\theta} \tag{6}$$

where the parameter θ is determined according to experiment (say $\theta = 2$).

We then proceed to construct the cost ρ_{i_f} due to flipping, which takes into account both the quantization and flipping effects if the original BMP cover image is available. Let x_i and q_i be the DCT coefficient and its quantization step, then $x\_q_i = x_i / q_i$ is the DCT coefficient after quantization. We further have the quantized DCT coefficient $\overline{x}_i$ and rounding error e_r_i

$$\begin{gathered} \overline{x}_i = round\left(x\_q_i\right) \\ e\_r_i = \overline{x}_i - x\_q_i \end{gathered} \tag{7}$$

The quantization error is then calculated as

$$e\_q_i = \overline{x}_i * q_i - x_i \tag{8}$$

For a binary embedding, the $\overline{x}_i$ should be inserted /removed by $+1/-1$ in the most appropriate positions to further decrease the flipping distortion. We then have the flipping rule, i.e., $\overline{x}_i \rightarrow y_i$

$$\begin{gathered} y_i = \overline{x}_i + 1 \quad if \;\; e\_r_i < 0 \\ y_i = \overline{x}_i - 1 \quad if \;\; e\_r_i \geq 0 \end{gathered} \tag{9}$$

The cost for flipping can then be defined as

$$\rho_{i\_f} = \left[\left(\left|y' - x\_q_i\right| - 0.5\right) * q_i\right]^2 \tag{10}$$

When the original BMP image is not available, the flipping cost in (10) can be simplified as

$$\rho_{i\_f} = q_i^{\,2} \tag{11}$$

Therefore, the proposed distortion function can be applied in JPEG steganography with/without original cover image support.

2.3. Encoder and decoder

The objective of of minimal distortion embedding framework is to improve embedding efficiency of steganography. To achieve this, syndrome coding based approach is usually employed.

The proposed framework for JPEG steganography utilizes the syndrome-trellis coding (STC) [1] for data hiding and the new distortion function for embedding efficiency optimization, which includes the encoder and decoder as shown in Fig.1.

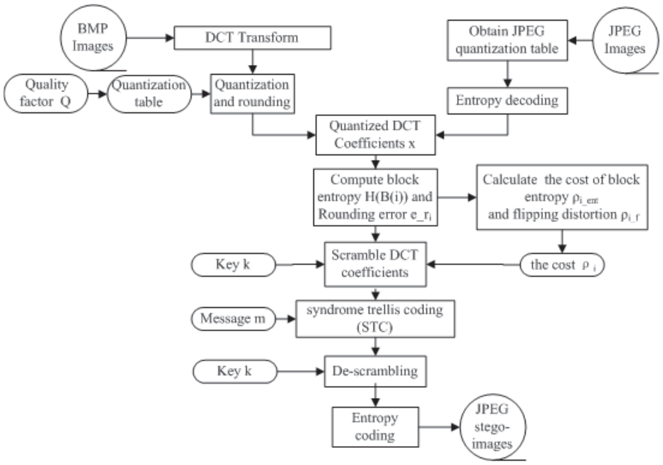
BMP Images
DCT Transform
Obtain JPEG quantization table
JPEG Images
Quality factor Q
Quantization table
Quantization and rounding
Entropy decoding
Quantized DCT Coefficients x
Compute block entropy H(B(i)) and Rounding error e_r$_i$
Calculate the cost of block entropy ρ_{i_ent} and flipping distortion ρ_{i_f}
Key k
Scramble DCT coefficients
the cost ρ_i
Message m
syndrome trellis coding (STC)
Key k
De-scrambling
Entropy coding
JPEG stego-images

Fig. 1. The proposed framework for JPEG steganography

The encoding process is described as follows:

1）Input

For BMP image input, the quantization table is determined according to the given QF Q . After the image is partitioned into 8×8 blocks and DCT transformed, the DCT coefficients are then quantized according to the given quantization table.

For JPEG image input, entropy decoding is applied to obtain the quantized DCT coefficients.

2) Compute costs ρ_{i_ent} and ρ_{i_f}

Compute the block entropy cost ρ_{i_ent} of 8×8 coefficients and the cost ρ_{i_f} due to flipping effect. And then determine the cost $\rho_i = \rho_{i\_ent} * \rho_{i\_f}$ for element i .

3) STC coding

Scramble the DCT coefficients with Key k to generate sequence **x** . The STC coding is then applied to embed message **m** to **x** to generate stego **y**

4) Stego JPEG image

After **y** is descrambled, the entropy coding is applied to generate stego JPEG image.

The decoding is the reverse process of encoding.

3 SIMULATION RESULTS AND ANALYSIS

Extensive experiments are carried out to verify the feasibility of the proposed JPEG steganographic scheme. The powerful steganalysis tool PEV274 developed by Pevny and Fridrich was used for experiments, which includes 193 DCT based features and 81 Markov features [2]. The 274 features from the cover and stego images are used to train the Support Vector Machine (SVM), which is obtained from LibSVM.

A set of 1,174 uncompressed gray scale images from CoreDraw database is used in our experiments. The test images includes pictures of different characteristics, such as landscapes, house, animals and plants, which are taken with cameras of different brand. The size of image is 768*512 or 512*768. We test the performance for different payloads (0.05 to 0.4 bpac) at the typical quality factor $QF = 75$

The error probability P_E (the minimum error probability under equal priors) is calculated as follows.

$$P_E = \min_{P_{FA}} \frac{1}{2}(P_{FA} + P_{MD}(P_{FA})) \qquad (12)$$

where the P_{FA} and P_{MD} denote the probability of false alarm and misdetection, respectively.

3.1 Visual Attack

To verify the visual indiscernibility between the cover and stego JPEG image with the proposed scheme, visual attacks are carried out. Fig.2 shows comparison of visual quality of cover JPEG images and their corresponding stego JPEG images of 15% bpac at QF=75. The top and bottom rows show cover and stego images, respectively. No obvious visual distortions due to embedding are observed in the stego images. (Fig. 2)

3.2 Statistical Attack

PEV-274 attack

Fig,3 shows the comparison in PE performance of the proposed scheme with other similar JPEG steganographic schemes, such as Jsteg and F5, STC with constant profile[1] and BCH+opt[11] . In our test, we use a STC code with $h = 10$ (the height of pseudo-random submatrix) for a better tradeoff between performance and complexity. The original BMP images are provided, and QF of the stego JPEG image is 75. It is observed that our scheme achieves significant improvement over other methods in performance against steganalysis with PEV274. For JPEG steganography at QF=75, the PE of PEV-274 attack remains above 40% even for payload $= 0.33$ bpac for the proposed scheme, i.e., the secure payloa $= 0.33$ bpac. Compared with the secure payload $= 0.17$ bpac achieved with BCH+Opt [11], which is the efficient JPEG steganographic scheme using BCH syndrome coding and heuristic optimization.

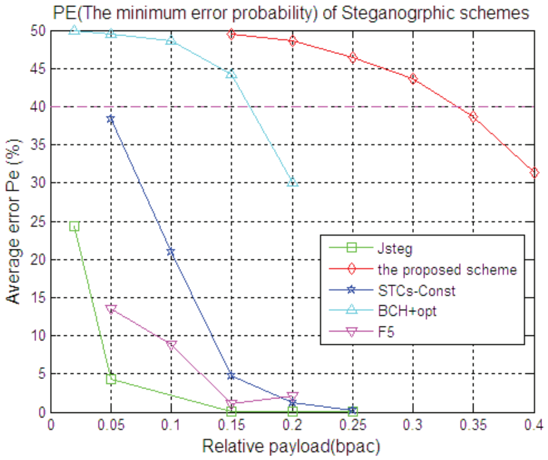

Fig. 3. The comparision of the proposed scheme with other schemes in different relative payloads

Other Universal Steganalysis Attack

We also test the performance of our scheme against other popular steganalysis tools besides PEV-274, i.e., Shi-78D [12] and MP-486[10]. The results are summarized in table I.

TABLE I
UNIVERSAL STEGANALYSIS ATTACK

Quantization Factor	**Relative Payload (bpac)**	**PE (*the minimum error probability under equal priors*)**		
		Pev274	Shi-78D	MP-486D
75	15%	49.50%	49.98%	50.08%
	25%	46.36%	50.02%	49.98%
	35%	38.60%	49.98%	50.24%
90	15%	49.92%	49.48%	50%
	25%	49.30%	50.08%	49.86%
	35%	44.68%	50.03%	47.46%

As we can see from Table I, the proposed scheme also performs well against Shi-78D and MP-486 analysises. Shi-78D is another common staganalysis tool, which is based on the moments of wavelet characteristic functions. While MP-486 utilizes both intra- and inter-block correlation among JPEG coefficients with Markov process [10], which shows good performance for staganalysis of steganographic schemes in DCT domain.

Steganography for JPEG format covers

The proposed scheme can also be used for JPEG steganography without original BMP image support, where the distortion function is characterized by the block entropy cost and a simplified flipping cost in (11). Although the performance of the scheme in secure payload is degraded significantly under the circumstances, it is much better than that of other popular scheme, such as F5. Fig.3 shows the performance of the proposed scheme without original BMP image support, where $QF = 90$

, $h = 10$ and $\theta = 4$. The proposed scheme can achieve a secure payload about 0.07 bpac ($PE \geq 40\%$), compared with the secure payload ≤ 0.02 bpac for F5.

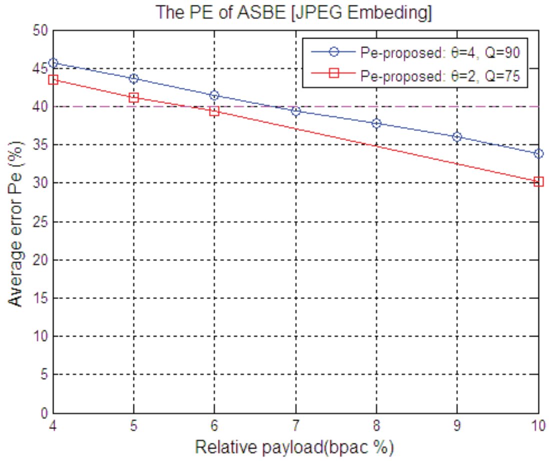

Fig. 4. The comparison of the proposed scheme with others in PE performance when the original BMP image is not available.

4 CONCLUSION

Minimal-distortion embedding framework is a practical approach to implement JPEG steganography with high embedding efficiency. This paper presents a new JPEG steganographic scheme which utilizes STC encoding based on a content adaptive distortion function. The STC provides multiple solutions to embed messages to a block of coefficients. And the steganographic scheme determines the best one with minimal distortion effect based on the distortion function. The proposed distortion function takes into account both the block entropy cost and the flipping cost, which guides the STCs to modify quantized DCT coefficients with minimal flipping distortion in regions of “hard-to-predict”, which leads to less detectability of steganalysis. Extensive experiments are carried out to demonstrate the superior performance of the proposed scheme in terms of secure embedding payload against steganalysis.

5 REFERENCES

[1] T. Filler, J. Judas, and J. Fridrich. “Minimizing Embedding Impact in steganography using Trellis-Coded Quantization,” *Proceedings SPIE, Electronic Imaging, Security and Forensics of Multimedia XII*, volume 7541, pages 05-01-05-14, San Jose, CA, pp. 17-21, Jan. 2010.

[2] T. Pevny and J. Fridrich. “Merging Markov and DCT Features for Multi-Class JPEG Steganalysis,”. Security, Steganography and Watermarking of Multimedia Contents IX, San Jose, CA, 6505, 2007

[3] A. Westfeld “F5-A Steganographic Algorithm: High Capacity Despite Better Steganalysis,” *4th Information Hiding International Workshop*, Pittsburgh, USA, Apr. 2001.

[4] Sallee, P.:”Model-Based Steganography”. Proc. of IWDW2003, pp. 154-167.

[5] D. Upham, JPEG–Jsteg. [Online].

[6] OutGuess - Practical Steganography [Online]

[7] J. Fridich and D. Soukual, “Matrix Embedding for Large Payload,” *IEEE Tran on Information Forensics and Security*, Vol 1, pp.390-295, Sep. 2006

[8] Y.H. Kim, Z Duric and D. Rechards, “Modified Matrix Encoding Technique for Minimal Distortion Steganography,” *Proc. Of Information Hiding 2006*, pp.314-327, 2007

[9] Y. Choi and H. Kim, “Improving the Modified Matrix Encoding on Steganography Method,” *Proc. IAS*,, 2009, pp.205-208.

[10] C. Chen and Y. Q. Shi, “JPEG image steganalysis utilizingboth intrablock and interblock correlations,” *IEEE International Symposium on Circuits and Systems*, Seattle, WA, pp.18-21, May. 2008.

[11] V. Sachnev, Hyoung-Joong Kim, Rongyue Zhang, “Less Detectable JPEG Steganography Method Based on Heuristic Optimization and BCH Syndrome Coding”. *Proceedings of the 11th ACM workshop on Multimedia and security*, pp.131~139, Sep. 2009.

[12] Y. Q. Shi, X. GUO, D ZOU, “Image Steganalysis Based on Moments of Characteristic Functions Using Wavelet Decomposition, Prediction-Error Image, and Neural Network,” *Proceedings of the IEEE International Conference on Multimedia and Expo*., pp.269~272, 2005.

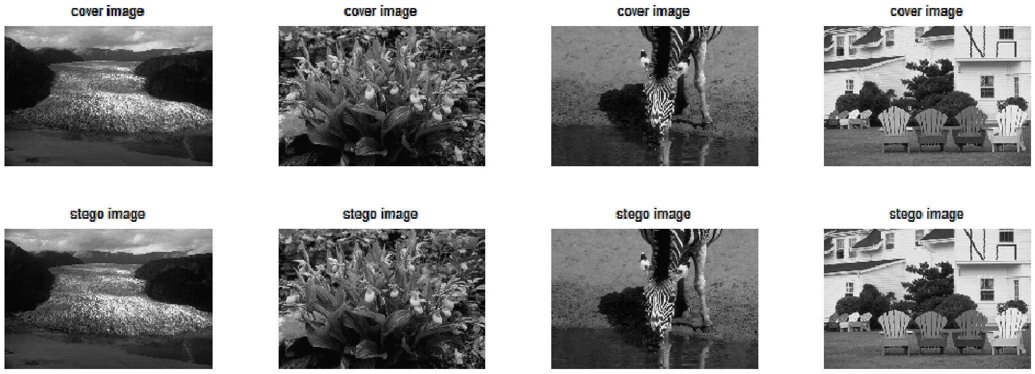

Fig.2 The comparison of visual difference between the cover and stego images with the proposed scheme